

**АКАДЕМИЯ НАУК СССР**  
**НАУЧНЫЙ ЦЕНТР БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ**  
**НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ВЫЧИСЛИТЕЛЬНЫЙ ЦЕНТР**  
**ИНСТИТУТ БЕЛКА**

**ПРЕПРИНТ**

**В.Ю.ЛУНИН, А.Г.УРЖУМЦЕВ, Е.А.ВЕРНОСЛОВА,**  
**Ю.Н.ЧИРГАДЗЕ, Н.П.ФОМЕНКОВА, Н.А.НЕВСКАЯ**

**ИСПОЛЬЗОВАНИЕ СМЕШАННОЙ МОДЕЛИ**  
**ЭЛЕКТРОННОЙ ПЛОТНОСТИ**  
**ДЛЯ УЛУЧШЕНИЯ НАБОРА ФАЗ**  
**В БЕЛКОВОЙ КРИСТАЛЛОГРАФИИ**

**ПУЩИНО • 1984**

УДК 548.73

Предложена методика уточнения значений фаз структурных факторов, использующая смешанную атомную модель распределения электронной плотности в исследуемом объекте. Эта модель состоит из атомов проинтерпретированной части молекулы (частичная модель) и набора фиктивных атомов, моделирующих электронную плотность в остальной части элементарной ячейки. Расчет фаз по смешанной модели, прошедшей уточнение в обратном пространстве, позволяет получить более надежные значения фаз, нежели имелись вначале. Это создает возможность построения более полной атомной модели молекулы.

Метод был применен при уточнении набора фаз и построении предварительной модели белка  $\lambda$ -кристаллина из глазной линзы телят при разрешении 2.7 Å. Стартовая частичная модель содержала координаты около 69% всех атомов. Было проведено два этапа уточнения, на каждом из которых получено значительное улучшение карт распределения электронной плотности, что позволило расширить частичную модель.

Предложенный метод является обобщением нескольких подходов к проблеме улучшения набора фаз, опубликованных ранее, и позволяет проводить построение атомных моделей белка при более низком разрешении, чем это было возможно ранее.

## 1. ВВЕДЕНИЕ

При определении пространственной структуры белка встречается ситуация, когда качество синтезов электронной плотности таково, что удается определить координаты лишь для части атомов молекулы белка. Построение же оставшейся части молекулы оказывается при этом невозможным без улучшения качества карт распределения электронной плотности. В данной работе излагается новый методический подход, использованный при работе с  $\gamma$ -кристаллином IIIb из глазной линзы телянка на этапе, когда была получена частичная модель молекулы и требовалось дальнейшее расширение модели.

В ситуации, когда отсутствует атомная модель, Агарвал и Айзекс /9/ предложили улучшать фазы путем уточнения грубой модели распределения электронной плотности. Под грубой моделью здесь понимается модель из "фиктивных" атомов, которая строится автоматически специальной программой и не претендует на какое-либо стереохимическое соответствие истинной структуре. Основное требование к грубой модели – удовлетворительно воспроизводить стартовое распределение электронной плотности в элементарной ячейке. Методика улучшения набора фаз путем построения и уточнения грубой модели была развита далее Луниным и Уржумцевым /4, 5/, опробована на актинидине в качестве теста и применена для  $\gamma$ -кристаллина при уточнении и расширении набора фаз с разрешением от 3.0 до 2.7 Å /7/.

В случае, когда частичная модель достаточно полна, то есть количество неопределенных атомов мало, для определения их положений могут быть использованы синтезы с экспериментальными значениями модулей структурных факторов  $F_o$  и фазами  $\varphi_o$ , рассчитанными по модели, а также разностные синтезы с коэффициентами типа

$$n F_o e^{i\varphi_o} - (n-1) F_c e^{i\varphi_c},$$

где  $F_o$  и  $\varphi_o$  – модули и фазы исходного синтеза,  $F_c$  и  $\varphi_c$  – модули и фазы, рассчитанные по модели,  $n$  – целое число.

Однако для успеха такого подхода требуется, чтобы частичная модель была достаточно полной и координаты ее атомов определены достаточно точно.

Райс показал, что значительного прогресса в повышении качества синтезов можно достичь, если подвергнуть частичную модель кристаллографическому уточнению в обратном пространстве и затем при построении синтезов использовать комбинацию фаз, рассчитанных по модели, и фаз, полученных методом изоморфного замещения /16/.

Бот и Блоу предложили другой способ использования частичной модели /10/. Вначале они сгенерировали электронную плотность, соответствующую интерпретированной части молекулы. Далее они расширили модель электронной плотности, включив в нее участки электронной плотности, примыкающие к интерпретированной части синтеза. Использование фаз, рассчитанных по такой расширенной модели электронной плотности, привело к повышению качества синтезов.

Излагаемая в данной работе методика использования частичной модели белка для улучшения набора фаз является обобщением всех изложенных выше подходов (табл. 1). Она заключается в следующем. На первом этапе мы создаем смешанную модель электронной плотности, дополняя частичную атомную модель молекулы фиктивными атомами так, чтобы воспроизвести с их помощью электронную плотность неинтерпретированной части синтеза. На втором этапе эта смешанная модель электронной плотности подвергается уточнению в обратном пространстве и рассчитанные по уточненной модели фазы используются для построения синтезов улучшенного качества.

Основное отличие нашего подхода от методики Райса состоит в том, что мы уточняем не просто частичную модель молекулы, а расширенную смешанную модель электронной плотности. Такая модель позволяет лучше аппроксимировать распределение электронной плотности в элементарной ячейке. Построение смешанной электронной плотности на первом этапе в принципе совпадает с предложенным в методике Бота и Блоу. Разница состоит в том, что Бот и Блоу строят смешанную модель плотности, допускающую произвольное распределение электронной плотности, а мы строим атомную смешанную модель, соответствующую этой электронной плотности. Тем самым мы вносим в дополнительную часть модели существенную информацию – атомность модели. Кроме того, мы не останавливаемся просто на пересчете новых значений фаз по смешанной модели, как в методике Бота и Блоу, а предварительно модифицируем и уточняем эту модель, добиваясь лучшего соответствия экспериментальным данным. Это позволяет надеяться на более высокое качество фаз, рассчитанных по модели.

Т а б л и ц а 1

Сопоставление различных подходов при уточнении фаз с использованием  
частичных моделей электронной плотности

Авторы	Предложенный метод	Интерпретированная часть электронной плотности (точечная атомная плотность)		Неинтерпретированная часть электронной плотности	
		использована	варьируется	использована	варьируется
Агарвал и Айзекс (1977)		нет	нет	да	да
Райс (1981)		да	да	нет	нет
Бот и Блоу (1982)		да	нет	да	нет
Лунин и др. (1984) (данная работа)		да	да	да	да

## 2. МЕТОДИКА ИСПОЛЬЗОВАНИЯ СМЕШАННОЙ МОДЕЛИ ЭЛЕКТРОННОЙ ПЛОТНОСТИ ДЛЯ УЛУЧШЕНИЯ НАБОРА ФАЗ

### 2.1. Исходная информация

Мы предполагаем, что в начале работы для некоторого набора рефлексов нам известны: а) экспериментально определенные значения модулей структурных факторов  $F_c$ ; б) распределения вероятностей  $P_o(\varphi)$  для соответствующих фаз структурных факторов, полученные на предыдущем этапе работы, например на основе метода изоморфного замещения. Тем самым мы можем определить "наилучшие" фазы  $\varphi_o$  и показатели достоверности  $m_o$  из равенства [12]

$$m_o e^{i\varphi_o} = \int_0^{2\pi} e^{i\varphi} P_o(\varphi) d\varphi;$$

в) координаты атомов  $\vec{z}_j$  и изотропные температурные факторы  $B_j$  частичной модели молекулы. Будем обозначать  $F_p e^{i\varphi_p}$  структурные факторы, рассчитанные по этой модели.

### 2.2. Построение смешанной модели электронной плотности

Смешанная модель, кроме атомов  $\{\vec{z}_j, B_j\}_{j=1}^N$  частичной модели молекулы, будет содержать еще некоторое количество фиктивных атомов с параметрами  $\{\vec{z}'_k, B'_k\}_{k=1}^M$ . Эти атомы выбираются таким образом, чтобы соответствующая их совокупности электронная плотность воспроизводила наиболее существенные детали разностного синтеза с коэффициентами

$$m_o (F_o e^{i\varphi_o} - k F_p e^{i\varphi_p}). \quad (1)$$

Введение коэффициента  $k$  вызвано тем, что в случае, когда частичная модель молекулы содержит сравнительно небольшую часть молекулы, средние значения  $F_p$  заметно меньше средних значений  $F_o$ . Это приводит к неполному вычитанию электронной плотности в местах нахождения атомов частичной модели. Поэтому, стремясь предотвратить попадание фиктивных атомов в близкую окрестность атомов частичной модели молекулы, мы ввели коэффициент  $k$ .

Определение параметров  $\vec{z}'_k, B'_k$  очередного фиктивного атома состояло из: а) поиска точки максимума  $\vec{z}'_k$  на разностном синтезе; б) определения параметра  $B'_k$  так, чтобы значение

электронной плотности в центре этого атома приближалось к значению плотности разностного синтеза в точке  $\vec{r}_k'$ ; в) вычитания вклада определенного фиктивного атома из синтеза. После этого происходил переход к построению следующего фиктивного атома до тех пор, пока максимальное значение "остаточной" электронной плотности не становилось меньше порогового значения  $\rho_{crit}$ . Более подробно процедура построения грубой модели неинтерпретированной части электронной плотности и определения вклада атома во взвешенный синтез описана ранее /7/.

На промежуточных этапах работы по расшифровке структуры белка может существовать дополнительная информация о расположении молекулы в элементарной ячейке. Например, могут быть известны границы молекулы. В таком случае естественно сузить область, из которой выбираются фиктивные атомы, разрешив точкам  $\vec{r}_k'$  лежать только внутри границ молекулы. Для того чтобы фиктивные атомы в процессе уточнения смешанной модели "не мешали" правильно сдвигаться атомам частичной модели, мы запрещали также фиктивным атомам находиться ближе, чем на заданном расстоянии от атомов частичной модели. Можно использовать и более тщательный выбор области, из которой берутся фиктивные атомы, например определять ее, следуя работе Бота и Блоу /10/, как связанные области, примыкающие к интерпретированной части синтеза.

### 2.3. Уточнение смешанной модели электронной плотности

Для уточнения смешанной модели мы использовали программу Айзекса, реализующую алгоритм Агарвала уточнения в обратном пространстве без учета стереохимических ограничений /8/. В то же время, по всей видимости, будет полезно на заключительных этапах работы вводить стереохимические ограничения на атомы частичной модели молекулы.

В процессе уточнения мы периодически производили просмотр атомов модели и удаляли из нее фиктивные атомы, для которых значение  $B_k'$  принимало достаточно большое значение.

### 2.4. Построение улучшенного синтеза

Для построения улучшенного синтеза использовались коэффициенты

$$m_{MR} F_0 e^{i\varphi_{MR}}, \quad (2)$$

где

$$m_{MR} e^{i\varphi_{MR}} = \int_0^{2\pi} e^{i\varphi} \rho_{MR}(\varphi) d\varphi \quad (3)$$

$$\rho_{MR}(\varphi) \sim e^{2tF_0} F_{MR} \cos(\varphi - \varphi_{MR}). \quad (4)$$

Здесь  $F_{MR} e^{i\varphi_{MR}}$  - структурные факторы, рассчитанные по уточненной смешанной модели электронной плотности;  $t$  - параметр, отражающий полноту и грубость модели, он был определен из максимума функции правдоподобия /3/.

Уточнение смешанной модели проводили по зоне  $d_{min} \leq d \leq d_{max}$ . Для рефлексов зоны  $d > d_{max}$  информацию о модели не использовали, а использовали исходное распределение вероятностей фаз  $\rho_0(\varphi)$ .

### 3. УЛУЧШЕНИЕ ФАЗ $\gamma$ -КРИСТАЛЛИНА IIIb ПРИ РАЗРЕШЕНИИ 2.7 Å

Кристаллины - основные структурные белки линзы позвоночных. Кристаллы  $\gamma$ -кристаллина IIIb с молекулярным весом около 20.0 кДа из глазной линзы телят имеют пространственную группу  $P2_12_12_1$  и размеры элементарной ячейки 58.7x69.5x116.9 Å /13/. В асимметричной части ячейки находятся две молекулы с общей массой около 40.0 кДа. Полипептидная цепь молекулы содержит около 170 остатков. Имеются предварительные данные по аминокислотной последовательности этого белка /15/, при этом полностью отсутствуют данные по участку цепи 37-55. Можно предполагать, что указанная последовательность содержит ряд ошибок. Это мнение основано на том, что определенная аналогичным химическим методом последовательность гомологичного белка  $\gamma$ -кристаллина II несколько отличается от таковой, определенной методом рентгеноструктурного анализа при разрешении около 2.6 Å /11, 14/.

Кристалл нативного белка  $\gamma$ -кристаллина IIIb имел дифракционное поле приблизительно до 2.7 Å. Ранее методом изоморфных замещений была получена карта электронной плотности с разрешением до 3.0 Å. Было использовано пять тяжелоатомных производных, три из них - с аномальным рассеянием, общее число отражений - 6777, средний фактор достоверности  $m = 0.72$  /13/. Фрагмент карты электронной плотности приведен в цитированной выше работе. Обе молекулы в асимметричной части ячейки были хорошо различимы на карте электронной плотности. В каждой из них выделялись два домена. Детально интерпретировали плотность только одной молекулы, поскольку распределение электронной плотности в обеих молекулах было очень близким. Четко был виден ход полипептидной цепи в областях вытянутых сегментов  $\beta$ -слоев ( $\gamma$ -кристаллин не содержит  $\alpha$ -спиралей), а также в области участка, соединяющего два домена между собой. Плохо просматривались некоторые участки петель молекулы



Т а б л и ц а 2

Улучшение фаз  $\lambda$ -кристаллина IIIb при разрешении 2.7 Å

Содержание	Этапы улучшения фаз		
	этап I непринтерпретированная электронная плотность	этап II использование частичной модели молекулы	этап III использование частичной модели молекулы
1	2	3	4

Стартовая модель

Количество атомов в час-  
точной атомной модели

нет

967

1130

Количество фиктивных атомов  
на 1/2 асимметричной части  
ячейки

1552

1444

873

Параметры\* конструирования  
модели из фиктивных атомов:

$B_0(\text{Å}^2)$ ,  $C_0(e)$ ,  $\rho_{crit}(e/\text{Å}^3)$

83.9, 8.40, 0.28

72.2, 9.11, 0.40

47.4, 7.69, 0.65

Учет грани молекулы\*

нет

да

да

Продолжение табл. 2

	1	2	3	4
Ход уточнения				
10,0-3,0 Å				
Количество циклов		10 (6хуз, 4В)	18 (12хуз, 6В)	31 (21хуз, 10В)
R <sub>initial</sub> → R <sub>final</sub>		0.35 → 0.13	0.38 → 0.14	0.46 → 0.16
10:0-2.7 Å				
Количество циклов		5 (3хуз, 2В)	18 (12хуз, 6В)	18 (12хуз, 6В)
R <sub>initial</sub> → R <sub>final</sub>		0.18 → 0.12	0.19 → 0.13	0.21 → 0.15
Характеристики полученного синтеза				
Число отражений		10527	10804	10804
Кoeffициенты синтеза		$\left\{ \begin{array}{l} m_0 F_0^{exp} i \varphi_0^{best} d > 10 \text{ \AA} \\ m_{comb} F_0^{exp} i \varphi_{comb}^{best} 10 \rightarrow d > 2.7 \text{ \AA} \end{array} \right.$	$\left\{ \begin{array}{l} m_0 F_0^{exp} i \varphi_0^{best} d > 10 \text{ \AA} \\ m_{comb} F_0^{exp} i \varphi_{comb}^{best} 10 \rightarrow d > 2.7 \text{ \AA} \end{array} \right.$	$\left\{ \begin{array}{l} m_0 F_0^{exp} i \varphi_0^{best} d > 10 \text{ \AA} \\ m_{comb} F_0^{exp} i \varphi_{comb}^{best} 10 \rightarrow d > 2.7 \text{ \AA} \end{array} \right.$
$\rho_{max}^{abs} (e/\text{Å}^3)$		2.06	2.83	3.07
$\rho_{min}^{map} (e/\text{Å}^3)$ *		0.56	0.55	0.46
(S/N) <sub>max</sub>		7.3	10.3	13.4

\* См. объяснение в тексте.

и нерегулярной конформации полипептидной цепи. Интерпретация боковых групп была весьма проблематичной. В результате качество электронной карты позволяло провести ход полипептидной цепи по  $C_{\alpha}$ -атомам с достаточной уверенностью приблизительно для 80% длины всей цепи. Предполагаемое число остатков при этом было равно 165.

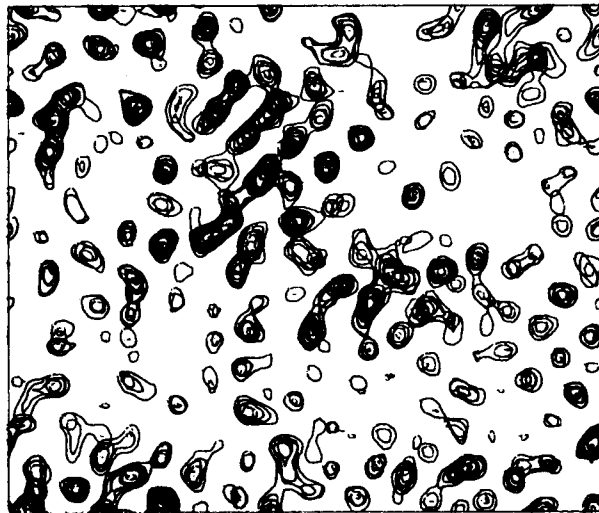
Описанная выше карта не позволяла интерпретировать боковые группы. Поэтому на первом этапе было проведено уточнение и расширение набора фаз в зоне 10–2.7 Å. Для этой цели была применена методика Уржумцева и Лунина /7/, реализующая идею Агарвала и Айзекса /9/. Сущность использованного на этом этапе подхода состояла в том, что в качестве модели для кристаллографического уточнения была взята грубая модель электронной плотности, состоящая из фиктивных атомов, воспроизводящих "изомерный" синтез. Какие-либо структурные данные не использовались.

Характеристики стартовой грубой модели и ход ее уточнения на первом этапе улучшения фаз описаны в табл. 2. В результате уточнения были дополнительно определены фазы около 3800 отражений, главным образом в зоне от 3.0 до 2.7 Å. Результирующий взвешенный синтез был построен на основе комбинированных фаз с использованием распределения вероятностей фазы вида

$$\rho_{comb}(\varphi) \sim \rho_o(\varphi) \rho_{MR}(\varphi).$$

Он представлен на рис. 1. По сравнению с изоморфным синтезом 3.0 Å новый синтез при 2.7 Å имеет более высокое максимальное значение плотности  $\rho_{max}^{abs} = 2.06 \text{ e/\AA}^3$  (было  $1.40 \text{ e/\AA}^3$ ), однако он более зашумлен. Тем не менее, новый синтез реально имеет лучшее разрешение. Его качество позволило приступить к построению частичной модели одной молекулы белка. При анализе электронной плотности был уточнен ход  $C_{\alpha}$ -цепи, общее предполагаемое число остатков возросло при этом до 167. С использованием приближенных данных по аминокислотной последовательности была проведена интерпретация электронной карты. Достаточно хорошо была локализована большая часть остатков Cys, Met, Phe, Tyr, которые обычно служат реперами при анализе электронных карт белков. Не удалось на этом этапе точно локализовать два остатка Trp, которые, по-видимому, находятся близко к местам изгибов основной цепи. Всего было вписано 93 боковые группы, то есть около 56% от предполагаемого общего числа. Частичная модель 1 молекулы содержала 967 атомов, то есть около 69% от общего их числа (табл. 3).

Частичная атомная модель молекулы была построена с помощью программы ISOCUB, использующей интерактивную молекулярную графику /6/. С помощью этой программы модель строит-



Z = 27 28 29 30

Рис. 1. Синтез  $\gamma$ -кристаллина при 2.7 Å, полученный после первого этапа уточнения набора фаз (без использования частичной модели). XY - проекция, Z = (27 ÷ 30)/144. В центре виден фрагмент одной из двух молекул асимметричной части ячейки

ся на основе приближенных координат  $C_\alpha$  и  $C_\beta$  атомов и вписывается в электронную плотность оператором последовательно для каждой пептидной единицы полипептидной цепи. Алгоритм позволяет быстро построить разные конформации этого фрагмента и найти наиболее приемлемое решение с точки зрения наилучшего вписывания его в электронную плотность и минимального искажения геометрии. Приближенное среднее отклонение центров тяжести отдельных групп от центров соответствующих электронных облаков составляло около 0.8 Å. Вся геометрия групп была близка к идеальной, за исключением углов  $\angle(NC_\alpha C)$ , соединяющих две пептидные группы, и угла  $\beta(C_\beta C_\alpha C)$ , соединяющего боковую группу и основную цепь.

На втором этапе улучшения фаз в качестве стартовой модели была использована смешанная модель электронной плотности. Она была составлена из 967 атомов частичной модели I и набора из 1444 фиктивных атомов, моделирующих неинтерпретированную область половины асимметричной части ячейки (табл. 2). При этом посадка фиктивных атомов допускалась только в пределах области, занятой молекулой. Объем последней определялся как совокупность сфер радиуса 5.5 Å, размещенных в центрах  $C_\alpha$ -атомов по всей цепи молекулы из 167 остатков. Вне этой области электронная плотность была принята равной нулю. Минимальное расстояние меж-

Т а б л и ц а 3

Характеристики частичных моделей молекулы  $\beta$ -кристаллина, использованные для улучшения фаз

Характеристики	Частичная модель	
	модель I	модель II
Предполагаемое общее число остатков в молекуле*	167	172
Предполагаемое общее число атомов в молекуле*	1411	1476
Число пептидных групп, в скобках - % к общему их числу	137 (82)	148 (86)
Число боковых групп, в скобках - % к общему их числу	93 (56)	118 (69)
Число атомов, в скобках - % к общему их числу	967 (69)	1130 (77)

Примечание. \* Точное число остатков в молекуле  $\beta$ -кристаллина IIIb неизвестно.

ду фиктивными атомами и атомами модели I на этом втором этапе улучшения фаз было равно  $0.7 \text{ \AA}$ . Число делений по осям элементарной ячейки при построении синтеза было выбрано равным  $100 \times 120 \times 200$ , что соответствует шагу приблизительно  $0.6 \text{ \AA}$  вдоль всех основных кристаллографических направлений. Посадку фиктивных атомов вели до исчерпания всей электронной плотности с  $\rho > \rho_{crit}$  на синтезе с коэффициентами (1). При этом коэффициент  $k$  (1) выбирался так, чтобы уравнивать средние значения  $F_o$  и  $k F_p$ . Вклад каждого фиктивного атома во взвешенный синтез моделировался функцией вида /5/

$$C_o \left( \frac{\pi}{B_o} \right)^{3/k} \exp \left( - \frac{4\pi^2 R^2}{B_o} \right). \quad (5)$$

Значения  $\rho_{crit}$ ,  $C_o$  и  $B_o$  указаны в табл. 2.

Таким образом была построена смешанная модель электронной плотности половины асимметричной части элементарной ячейки. Вторая половина асимметричной части заполнялась с использованием некристаллографической симметрии.

Уточнение смешанной модели электронной плотности проводили двумя сериями (табл. 2). В первой серии из 18 циклов уточнение проводили для зоны  $10-3.0 \text{ \AA}$ , содержащей около 8600 структурных факторов. Чередовали уточнение по координатам ( $x, y, z$ ) и температурным факторам B-атомов. Кристаллографический фактор

$$R = \frac{\sum |F_o - F_m|}{\sum F_o}$$

для стартовой модели был равен 0.38 и снизился в результате уточнения до 0.14. Уточнение во второй серии проводили в зоне  $10-2.7 \text{ \AA}$ , содержащей около 10800 структурных факторов, то есть в полтора раза больше, чем в первой серии. В итоге получена величина  $R = 0.13$ . Следует иметь в виду, что финальное значение этой величины гораздо ниже такового при обычном уточнении, когда уточняют не смешанную, а полную атомную модель молекулы белка, что связано с большим числом степеней свободы смешанной модели.

Полученный на втором этапе синтез имел значительно лучшее качество (рис. 2). Сравнение некоторых характеристик полученных синтезов после первого и второго этапов можно получить из данных табл. 2. Значительно возрос контраст электронной карты. Так, значение  $\rho_{max}^{abs}$  возросло в 1.35 раза. Также возросло отношение сигнал-шум. Его для синтеза, вычисленного в абсолютной шкале, мы предлагаем оценивать величиной

$$(S/N)_{max} = 2 \left( \rho_{max}^{abs} / \rho_{min}^{map} \right). \quad (6)$$

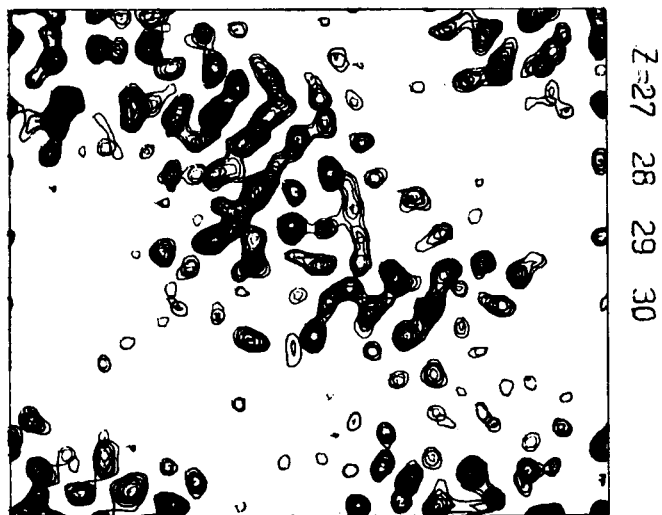


Рис. 2. Синтез  $\gamma$ -кристаллина при  $2.7 \text{ \AA}$ , полученный после второго этапа уточнения набора фаз по методу, предложенному в данной работе. Вывод проекции, как на рис. 1. Детали вывода синтеза в табл. 2

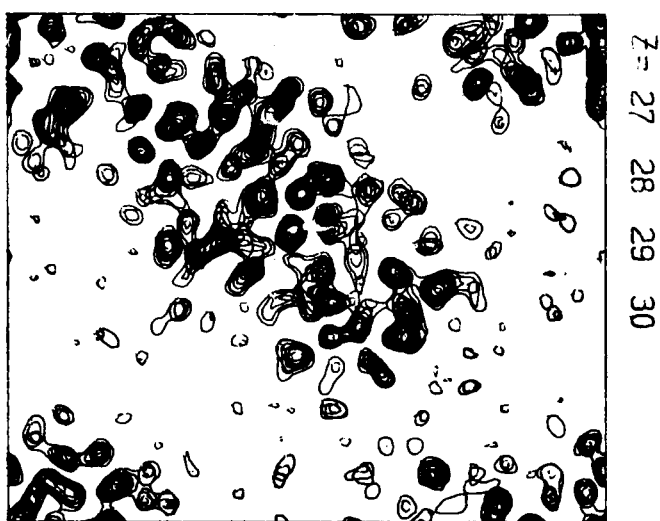


Рис. 3. Синтез  $\gamma$ -кристаллина при  $2.7 \text{ \AA}$ , полученный после третьего этапа уточнения набора фаз по методу, предложенному в данной работе. Вывод проекции, как на рис. 1. Детали вывода синтеза в табл. 2

Здесь  $\rho_{max}^{abs}$  - значение максимума электронной плотности,  
 $\rho_{min}^{map}$  - значение первого уровня электронной плотности, который  
 прорисовывается на карте при подготовке синтеза для анализа.  
 Значение  $\rho_{min}^{map}$  выбирается из субъективных соображений, одна-  
 ко в любом случае оно выше уровня шума. Мы будем полагать,  
 что уровень шума равен величине  $\rho_{min}^{map}/2$ .

После проведения второго этапа уточнения значение  $(S/N)_{max}$   
 увеличилось в 1,38 раза. Синтез стал значительно яснее, лучше  
 прорисовывался ход основной цепи. В некоторых местах на элек-  
 тронной плотности основной цепи появились "вздутия", указываю-  
 щие на локализацию пептидного карбонила. Существенно лучше  
 стали интерпретироваться электронные облака боковых групп. Фак-  
 тически уже по результатам этого этапа можно судить об эффек-  
 тивности использованного нами метода улучшения фаз. Мы смог-  
 ли приступить к построению предварительной атомной модели бел-  
 ковой молекулы при сравнительно низком разрешении и недоста-  
 точно надежных данных по аминокислотной последовательности.

Синтез, полученный на втором этапе, позволил уточнить и  
 значительно расширить модель молекулы белка. Был уточнен ход  
 основной цепи, предполагаемое общее количество остатков стало  
 равным 172. В электронную плотность было вписано 118 боко-  
 вых групп, то есть около 69% от их общего числа. Частичная мо-  
 дель II содержала уже 1130 атомов, что составило около 77%  
 от их общего числа. Лучше стала статистика геометрических ис-  
 кажений по деформируемым при построении модели углам  $\alpha$  и  $\beta$ .

Смешанная модель для третьего этапа была составлена из  
 1130 атомов частичной модели II и набора 873 фиктивных ато-  
 ма, моделирующих электронную плотность в неинтерпретированной  
 области половины асимметричной части ячейки (табл. 2). После  
 этого для каждого атома в смешанную модель был добавлен атом,  
 связанный с ним операцией некристаллографической симметрии.  
 Уточнение было проведено 49 циклами в двух диапазонах разре-  
 шения. Окончательное значение R-фактора равно 0,15. Получен-  
 ный в результате синтез был лучшего качества, чем в результа-  
 те второго этапа улучшения фаз (рис. 3). Возрос контраст карты:  
 значение  $\rho_{max}^{abs}$  увеличилось в 1,09 раза, а отношение сигнал-  
 шум  $(S/N)_{max}$  - в 1,30 раза. Качество синтеза позволило  
 провести дальнейшее уточнение модели и увеличить ее полноту.

#### 4. ОБСУЖДЕНИЕ

Применение предложенного метода к уточнению фаз  $\gamma$ -крис-  
 таллина IIIb из глазной линзы телянка при разрешении 2,7 Å  
 показало его высокую эффективность. На каждом из двух этапов



уточнения значительно улучшился синтез электронной плотности. Так, если использовать введенную выше оценку отношения сигнал-шум, то эта величина возросла в итоге почти в два раза. Значителен и прогресс в расширении частичной атомной модели молекулы белка: ее полнота возросла с 69% до 80%. Практически стал возможным переход к обычному последнему этапу анализа – уточнению структуры молекулы белка с использованием стереохимических ограничений.

Предложенная методика позволяет на протяжении всего процесса построения атомной модели объекта использовать наличие уже проинтерпретированных участков синтеза для улучшения качества остальной его части. При этом в начале работы, когда количество включенных в модель атомов невелико, смешанная модель состоит в основном из фиктивных атомов, и методика работы идейно близка к методу Агарвала и Айзекса. На заключительных же стадиях, когда относительное количество фиктивных атомов становится мало, уточнение смешанной модели переходит в обычное кристаллографическое уточнение атомной модели в обратном пространстве.

Использование фиктивных атомов на промежуточных стадиях работы (когда частичная модель содержит около половины всех атомов) позволяет для рассчитанных по модели фаз снизить уровень ошибок, вызванных отсутствием части молекулы. При уточнении смешанной модели без учета стереохимических ограничений наличие фиктивных атомов играет также стабилизирующую роль, не давая атомам частичной модели уйти далеко от правильных положений. Заметим, что использование стереохимических ограничений для атомов частичной модели на заключительных этапах работы по улучшению качества фаз может, по-видимому, привести к дальнейшему прогрессу.

Наш опыт использования грубой и смешанной моделей для улучшения набора фаз ограничивается зоной разрешения 3,5–2,0 Å. При работе с более высоким разрешением, как правило, имеется достаточно полная атомная модель и потребность введения в нее фиктивных атомов невелика. Однако и здесь эта методика может оказаться полезной, например для уточнения распределения электронной плотности в межмолекулярной области. При низком разрешении сходная идея была использована на практике для белков типа миоглобина при разрешении 5 Å /1, 2/. Здесь грубая модель строилась из "квазицилиндров", имитирующих  $\alpha$ -спирали, параметры которых уточнялись. Возможности использования обсуждаемой методики в области 5–3,5 Å в настоящее время не изучены. Здесь, по-видимому, грубую модель нужно строить из каких-то более крупных блоков, нежели атомы, чтобы добиться разумного соотношения между числом параметров, описывающих модель, и числом используемых в процессе уточнения структурных факторов.

В современной белковой кристаллографии можно выделить сле-

дующие необходимые основные этапы: 1) определение фаз методом изоморфного замещения; 2) уточнение фаз и построение предварительной модели белковой молекулы; 3) уточнение окончательной пространственной атомной модели молекулы белка.

В большинстве случаев для белков среднего и большого молекулярного веса методом изоморфного замещения не удастся получить фазы, которые можно было бы без уточнения использовать для построения приемлемой модели белковой молекулы. Именно поэтому необходимо использовать различные способы введения дополнительной информации для улучшения набора фаз. Одним из возможных вариантов рассмотрения этой задачи может быть подход, предложенный нами в данной работе.

#### ЛИТЕРАТУРА

1. Боровиков В.А., Вайнштейн Б.К., Гельфанд И.М., Калинин Д.И. - Кристаллография, 1979, т. 24, с. 227-238.
2. Калинин Д.И. - Кристаллография, 1980, т. 25, с. 535-544.
3. Лунин В.Ю. Использование метода максимального правдоподобия для оценки ошибок при определении фаз в кристаллографии белка. Препринт. Пушино, ОНТИ НЦБИ АН СССР, 1982.
4. Лунин В.Ю., Уржумцев А.Г. Повышение разрешения карт электронной плотности белков путем уточнения модельной структуры. I. Описание метода. Препринт. Пушино, ОНТИ НЦБИ АН СССР, 1981.
5. Лунин В.Ю., Уржумцев А.Г. - Кристаллография, 1983, т. 28, с. 310-315.
6. Невская Н.А., Курочкина Н.А., Чиргадзе Ю.Н. Построение пространственной модели белка по картам электронной плотности с использованием молекулярной графики. Препринт. Пушино, ОНТИ НЦБИ АН СССР, 1983.
7. Уржумцев А.Г., Лунин В.Ю. Повышение разрешения карт электронной плотности белков путем уточнения модельной структуры. II. Улучшение набора фаз для  $\gamma$ -кристаллина IIIb. Препринт. Пушино, ОНТИ НЦБИ АН СССР, 1982.
8. Agarwal R.C. - Acta Cryst., 1978, v. A34, p. 791-809.
9. Agarwal R.C., Isaacs N.W. - Proc. Natl Acad. Sci. USA, 1977, v. 74, p. 2835-2839.
10. Bhat T.N., Blow D.M. - Acta Cryst., 1982, v. A38, p. 21-29.
11. Blundell T.L. et al. - Nature, 1981, v. 289, p. 771-777.
12. Blow D.M., Crick F.H.C. - Acta Cryst., 1959, v. 12, p. 794-802.

13. Chirgadze Yu.N. et al. - FEBS Lett., 1981, v. 131, p. 81-84.
14. Croft L.R. - Biochem. J., 1972, v. 128, p. 961-970.
15. Croft L.R., Slingsby G. - In: The Human Lens - In Relation to Cataract, Ciba Found. Symp. v. 19, Amsterdam-N.Y., Elsevier-Excerpta Medica, 1973, p. 212-213.
16. Rice D.W. - Acta Cryst., 1981, v. A37, p. 491-500.

Владимир Юрьевич Лунин, Александр Георгиевич Уржумцев,  
Елена Анатольевна Вернослава, Юрий Николаевич Чиргадзе,  
Наталья Павловна Фоменкова, Наталья Александровна Невская  
ИСПОЛЬЗОВАНИЕ СМЕШАННОЙ МОДЕЛИ ЭЛЕКТРОННОЙ  
ПЛОТНОСТИ ДЛЯ УЛУЧШЕНИЯ НАБОРА ФАЗ  
В БЕЛКОВОЙ КРИСТАЛЛОГРАФИИ

Препринт

Отредактировано и подготовлено к печати в ОНТИ НЦБИ

T23318. Подписано в печать 30.11.83 г.

Уч.-изд.л. 1,0. Тираж 200 экз. Заказ 4167Р.

Бесплатно. Изд. № 474.

Отпечатано на ротапринте в Отделе научно-технической информации  
Научного центра биологических исследований АН СССР в Пушкине